Exon IB is primarily considered herein to encode the indicated 31 amino acids. However, the splice occurs within a codon for aspartate. This aspartate residue is common to both isoforms of utrophin. In embodiments of the invention an aspartate residue may be included C-terminal to the 31 amino acids to provide a 32 amino acid peptide, which may be joined to additional amino acids, for instance additional utrophin sequence as discussed. See, for instance, Figure 8 (SEQ ID NO: 7) for one embodiment.

These findings significantly contribute to the understanding of the molecular physiology of utrophin expression and are important because the promoter reported here provides an alternative target for transcriptional activation of utrophin in DMD muscle. This promoter does not contain synaptic regulatory elements and might, therefore, be a more suitable target for pharmacological manipulation than the previously described promoter.

We have now cloned this alternative utrophin promoter and exon, and the present invention in various aspects and embodiments is based on the sequence information obtained and provided herein.

One major use of the promoter is in screening for substances able to modulate its activity. It is well known that pharmaceutical research leading to the identification of a new drug generally involves the screening of very large numbers of candidate substances, both before and even after a lead compound has been found. This is one factor which makes pharmaceutical research very expensive and time-consuming. A method or means assisting in the screening process will have considerable commercial importance and utility. Substances identified as upregulators of the utrophin promoter represent an advance in the fight against muscular dystrophy since they provide basis for design and investigation of therapeutics for

in vivo use.

In one aspect, the present invention provides an isolated nucleic acid comprising a promoter, the promoter comprising a sequence of nucleotides shown in Figure 1 (SEQ ID NO: 1) or Figure 2 (SEQ ID NO: 3). The promoter may comprise one or more fragments of the sequence shown in Figure 1 of Figure 2 sufficient to promote gene expression. The promoter may comprise or consist essentially of a sequence of nucleotides 5' to position 1440 in Figure 1 (human) or position 1183 in Figure 2 (mouse). Preferably the promoter comprises or consists essentially of nucleotides 1199 to 1440 of the human sequence shown in Figure 1, or the equivalent sequence in mouse, e.g. nucleotides 959 to 1183 of Figure 2.

An even smaller portion of this part of the sequences shown in Figure 1 of Figure 2 may be used as long as promoter activity is retained. Restriction enzymes or nucleases may be used to digest the nucleic acid, followed by an appropriate assay (for example as illustrated herein using luciferase constructs) to determine the minimal sequence required. A preferred embodiment of the present invention provides a nucleic acid isolate with the minimal nucleotide sequence shown in Figure 1 or Figure 2 required for promoter activity. The minimal promoter element is situated between the PvuII restriction site at position 1199 in the human sequence and the transcription start site at 1440 bp in the human sequence and between nucleotides 959 to 1183 in the mouse sequence (see Figure 2).

In one embodiment a promoter according to the present invention comprises or consists of sequence that is shown in Figure 3 to be conserved between the human and mouse sequences, e.g. the 25 nucleotide sequence: ACAGGACATCCCAGTGTGCAGTTCG (SEQ. ID NO: 10) spanning the transcriptional start site.

expression.

Following identification of a substance which modulates or affects utrophin promoter activity, the substance may be investigated further. Furthermore, it may be manufactured and/or used in preparation, i.e. manufacture or formulation, of a composition such as a medicament, pharmaceutical composition or drug. These may be administered to individuals.

As noted above, the inventors also identified a novel coding sequence (Exon IB) which encodes a novel utrophin N-terminus.

According to a further aspect of the present invention there is provided a nucleic acid molecule which has a nucleotide sequence encoding a polypeptide which includes the amino acid sequence shown in Figure 1 (SEQ ID NO: 2) or Figure 2 (SEQ ID NO: 4) Such a polypeptide may include other utrophin sequences, and the nucleic acid molecule may be in the form of a utrophin "mini-gene" (discussed further below).

Such a polypeptide may include non-utrophin (i.e. heterologous or foreign) sequences and thereby form a larger fusion protein. For example, such a fusion protein could be used to target a non-utrophin polypeptide to muscle membranes.

The coding sequence included may be that shown in Figure 1 or Figure 2 or it may be a mutant, variant, derivative or allele of the sequence shown. The sequence may differ from that shown by a change which is one or more of addition, insertion, deletion and substitution of one or more nucleotides of the sequence shown. Changes to a nucleotide sequence may result in an amino acid change at the protein level, or not, as determined by the genetic code.

Thus, nucleic acid according to the present invention may

that the polypeptide comprises an actin-binding domain and a DPC-binding domain but is shorter than naturally occurring utrophin. In a full-length utrophin gene including what are identified herein as exons 1A and 1B, the actin-binding domain is encoded by nucleotides 1-739, while the DPC-binding domain (CRCT) is encoded by nucleotides 8499-10301 (where 1 represents the start of translation). See also Figure 8 (SEQ ID NO. 5) The respective domains in the polypeptide encoded by a mini-gene according to the invention may comprise amino acids corresponding to those encoded by these nucleotides in the full-length coding sequence. In one embodiment, a minigene according to the present invention comprises or consists of the amino acid sequence encoded by nucleotides 1-739 and 8499-10301 of the A isoform of utrophin in which exon 1B as identified herein is substituted for exons 1A and 2A. The sequence of such a minigene can be constructed by the ordinary skilled person using information disclosed herein, taking into account the content of WO97/22696 and Tinsley et al, *Nature* (1996) 384:349. The nucleic acid sequence and predicted amino acid sequence encoded by a `mini-gene' according to the present invention are shown in Figure 9 (SEQ ID NO: 8)

Advantages of a mini-gene over a sequence encoding a full-length utrophin molecule or derivative thereof include easier manipulation and inclusion in vectors, such as adenoviral and retroviral vectors for delivery and expression.

A further preferred non-naturally occurring nucleic acid molecule encoding a polypeptide with the specified characteristics is a chimaeric construct wherein the encoding sequence comprises a sequence obtainable from one mammal, preferably human ("a human sequence"), and a sequence obtainable from another mammal, preferably mouse ("a mouse sequence"). Such a chimaeric construct may of course comprise the addition, insertion, substitution and/or deletion of one or more nucleotides with respect to the parent mammalian

sequences from which it is derived. Preferably, the part of the coding sequence which encodes the actin-binding domain comprises a sequence of nucleotides obtainable from the mouse, or other non-human mammal, or a sequence of nucleotides derived from a sequence obtainable from the mouse, or other non-human mammal.

In a preferred embodiment, the sequence of nucleotides encoding the polypeptide comprises sequence GAGGCAC at residues 331-337 and/or the sequence GATTGTGGATGAAAACAGTGGG (SEQ. ID No: 11) at residues 1453-1475 (using the conventional numbering from the initiation codon ATG), and a sequence obtainable from a human.

Nucleic acid according to the present invention is obtainable using one or more oligonucleotide probes or primers designed to hybridise with one or more fragments of a nucleic acid sequence shown in Figure 1 or Figure 2 particularly fragments of relatively rare sequence, based on codon usage or statistical analysis. The amino acid sequence information provided may be used in design of degenerate probes/primers or "long" probes. A primer designed to hybridise with a fragment of the nucleic acid sequence shown may be used in conjunction with one or more oligonucleotides designed to hybridise to a sequence in a cloning vector within which target nucleic acid has been cloned, or in so-called "RACE" (rapid amplification of cDNA ends) in which cDNA's in a library are ligated to an oligonucleotide linker and PCR is performed using a primer which hybridises with the sequence shown in the figures and a primer which hybridises to the oligonucleotide linker.

Nucleic acid isolated and/or purified from one or more cells (e.g. human, mouse) or a nucleic acid library derived from nucleic acid isolated and/or purified from cells (e.g. a cDNA library derived from mRNA isolated from the cells), may be probed under conditions for selective hybridisation and/or subjected to a specific nucleic acid amplification reaction

between energy states result in characteristic spectral absorptions or emissions. They may include chemical entities used in conjunction with biosensors. Biotin/avidin or biotin/streptavidin and alkaline phosphatase detection systems may be employed.

The mode of determining binding is not a feature of the present invention and those skilled in the art are able to choose a suitable mode according to their preference and general knowledge. Particular embodiments of antibodies according to the present invention include antibodies able to bind and/or which bind specifically, e.g. with an affinity of at least 10^{-7} M, to the peptides shown in Figure 1 (SEQ ID NO: 2) or Figure 2 (SEQ ID NO: 4).

Antibodies according to the present invention may be used in screening for the presence of a polypeptide, for example in a test sample containing cells or cell lysate as discussed, and may be used in purifying and/or isolating a polypeptide according to the present invention, for instance following production of the polypeptide by expression from encoding nucleic acid therefor.

An antibody may be provided in a kit, which may include instructions for use of the antibody, e.g. in determining the presence of a particular substance in a test sample. One or more other reagents may be included, such as labelling molecules, buffer solutions, elutants and so on. Reagents may be provided within containers which protect them from the external environment, such as a sealed vial.

The present invention extends in various aspects not only to a substance identified using a nucleic acid molecule as a modulator of utrophin promoter activity, or to a polypeptide, or nucleic acid molecule in accordance with what is disclosed herein, but also a pharmaceutical composition, medicament, drug or other composition comprising such a substance, a

utrophin, showing a nucleotide sequence and encoded polypeptide according to embodiments of the present invention.

Figure 9 shows the nucleic acid and predicted amino acid sequence of a utrophin B isoform `minigene'.

Figure 10 shows the dosage dependence of IL-6 mediated expression from the isoform B promoter.

Oligonucleotides, PCR, RT-PCR and 5'RACE

PCR and RT-PCR were performed as described (Blake, *et al.* (1996) *J Biol Chem* 271, 7802-7810). Oligonucleotide sequences (5' to 3') were:

UM83	gatgttcctg tgaggccttc gag,	(SEQ ID NO: 12)
UM82	cactcttgga aaatcgagcg t,	(SEQ ID NO: 13)
U16	actatgatgt ctgccagagt tg,	(SEQ ID NO: 14)
U107	gatccaatag cttccttcca tcttt,	(SEQ ID NO: 15)
UBF	tggaaaaagt ggaggttgga,	(SEQ ID NO: 16)
BR2	tccaacctcc actttttcca,	(SEQ ID NO: 17)
BR4	gcctggagag ctacatgccc t,	(SEQ ID NO: 18)
BF8	ctccacatct ttttcctcat catct,	(SEQ ID NO: 19)
BF9	gattgtggtg atggttgtag aa,	(SEQ ID NO: 20)
BR10	gattgtggtg atggttgtag aa,	(SEQ ID NO: 20)
BR14	gatgatgagg aaaaagatgt ggag,	(SEQ ID NO: 21)
BF15	aaacccaaaa taacacagga catc,	(SEQ ID NO: 22)
BF16	agtgtaactt ctctctggtg,	(SEQ ID NO: 23)
BF31	taagcagatg taggtgatga gc,	(SEQ ID NO: 24)
BF42	gctgcttttg ttgtccactt c,	(SEQ ID NO: 25)
BR43	atagcttcct tccatctttg ag,	(SEQ ID NO: 26)
CT2	ctccacgttc ttccctctct act,	(SEQ ID NO: 27)
2ApF	gcgtgcagtg gaccattttt cagattta,	(SEQ ID NO: 28)
1BpF	cgctgcagca gccaccacat ttcgttg,	(SEQ ID NO: 29)
3pR	gcgtgcagat cgagcgttta tccatttg.	(SEQ ID NO: 30)

5' RACE was undertaken using adapter-ligated mouse heart cDNA (Marathon-Ready, Clontech), following the manufacturer's